Thesis/
Reports
Mead,
R. A.

DEVELOPMENT OF AN AERIAL PHOTO VOLUME TABLE FOR TREES INFESTED WITH SOUTHERN PINE BEETLES

Roy A. Mead
James L. Smith

FINAL REPORT

No. 80-1

Development of an Aerial Photo Volume Table for Trees Infested with Southern Pine Beetles

by

Roy A. Mead[1]
James L. Smith

November, 1979

School of Forestry and Wildlife Resources
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

Project supported by
Forest Insect and Disease Management
U. S. Forest Service
State and Private Forestry
Atlanta, Georgia

Cooperative Agreement Number 42-282

FINAL REPORT

No. 80-1

Development of an Aerial Photo Volume Table for Trees Infested with Southern Pine Beetles

by

Roy A. Mead[1]
James L. Smith

November, 1979

[1]The authors are Assistant Professor and Graduate Research Assistant, respectively, in the School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061.

LIBRARY COPY
ROCKY MTN. FOREST & RANGE
EXPERIMENT STATION

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Aerial photo volume tables have been developed for several timber types over several regions of the United States. These tables predict timber volume with varying degrees of reliability and are subject to local site conditions. Therefore, existing volume tables must be evaluated for suitability and precision before they can be implemented on an operational basis in any specific region. In some cases existing tables can be slightly modified but often new tables must be developed. In this study a new aerial photo stand volume table was developed for a specific area and compared with the existing tables for the region which includes the study area.

A. The specific objectives were:

1. To evaluate existing tree volume tables and stand volume tables for use in the study area on trees (or stands) infested with southern pine beetles.
2. To develop new tree volume tables and stand volume tables for the study area on trees (or stands) infested with southern pine beetles.
3. To evaluate the reliability of the new tables and recommend to the Forest Service which tables (old or new) should be used.
4. Identify the variables which influence the reliability of the photo/volume estimation methodology.

B. Justification:

Certainly an efficient, inexpensive, and timely means of assessing the volume of timber killed by the southern pine beetle needs to be developed. Aerial photo stand volume tables, when used with a photo sampling methodology, would permit a photo-interpreter to estimate timber volume loss for stands infested by the southern pine beetle.

C. Study Area:

A three million acre area in central Mississippi including National Forest, timber industry, state and small private ownership, was chosen for this study (Figure 1). It included conditions representative of SPB outbreak areas throughout most of the south. Both high and low levels of beetle-caused tree mortality occurred within the area. The study area was relatively flat, thus keeping scale deviation due to relief displacement at a minimum. Considerable variation in pine age, height and degree of mixture with hardwood occurred in the area. Both plantations and natural stands were present in the selected site.

DEVELOPMENT OF THE AERIAL PHOTO STAND VOLUME TABLE

A. Model Variables:

The objective of this study was to construct an aerial photo stand volume table utilizing measurements taken on 1:8,000 scale color infrared aerial photographs. While some investigators have used photo

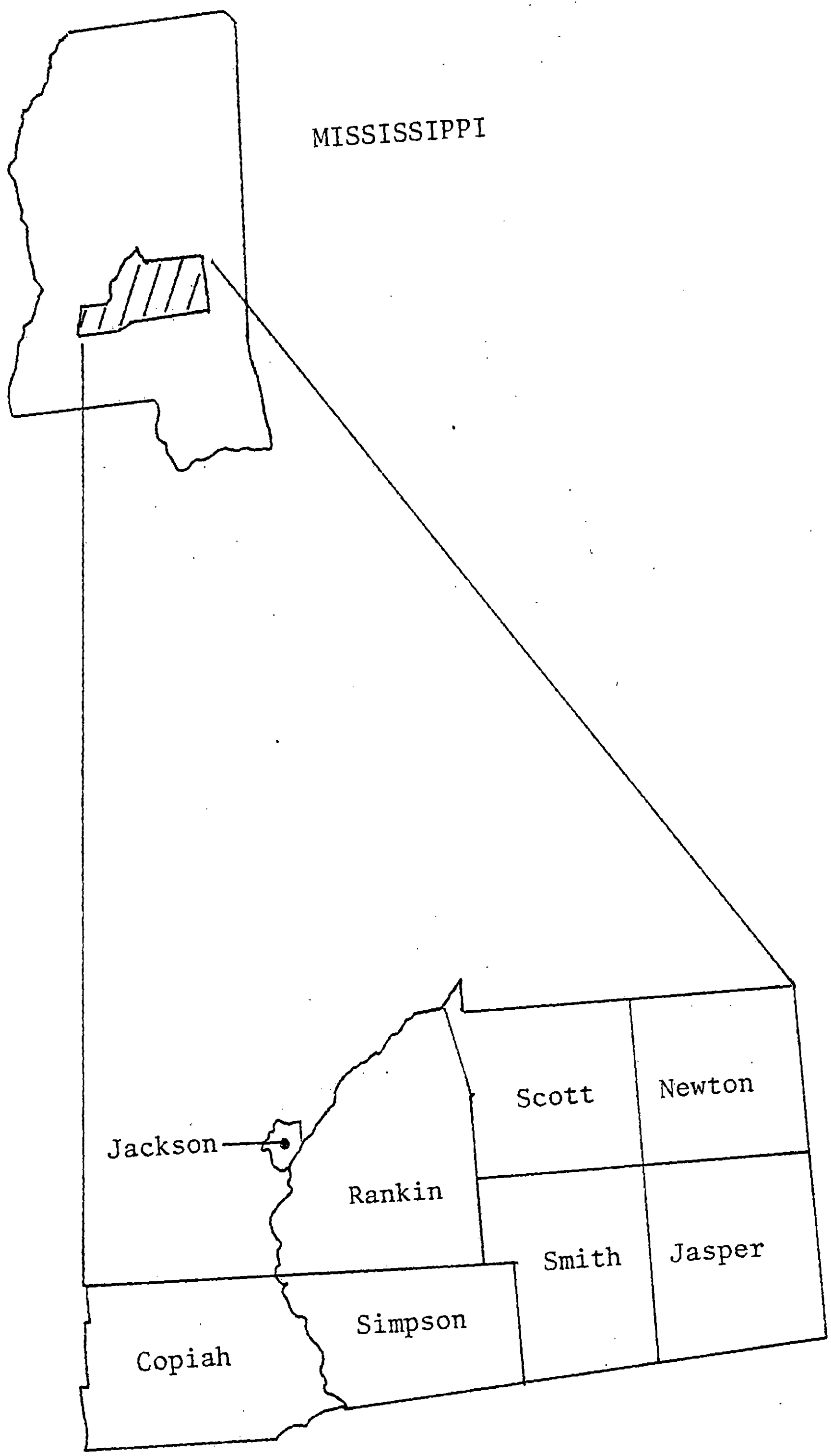

Figure 1. Map of the study area.

measures of crown diameter and estimates of the number of trees per unit area for predicting volume, this was not done in this study. It was concluded that due to the site conditions, scale and season of the photography, only estimates of percent crown closure of pine and average tree height would be used to predict volume.

Crown closure was defined as the percent of ground area covered by live tree crowns, i.e., an indicator of site occupancy. In this study, the range of crown closure was partitioned into 10 disjoint classes of width 10%. The value given to each class is its midpoint. Five different interpreters estimated crown closure on each of 50 fifth-acre circular plots (Figure 2). The class enclosing the average of the 5 measurements was taken as the actual value of crown closure for that plot.

For this investigation, height was defined as the average total height of the four tallest trees within a fifth acre plot. For development of the volume table, the ground estimate of height was used. However, when applying the table to other stands, height would be estimated from the photographs. Estimates of tree heights from aerial photographs requires measuring the parallax difference, photographic scale and the air base. Measurement errors and/or bias in any of these three variables could have affected the selection of a proper volume prediction model. Therefore, the investigators felt justified in using the field estimated tree heights for constructing the aerial photo stand volume table (Figure 3).

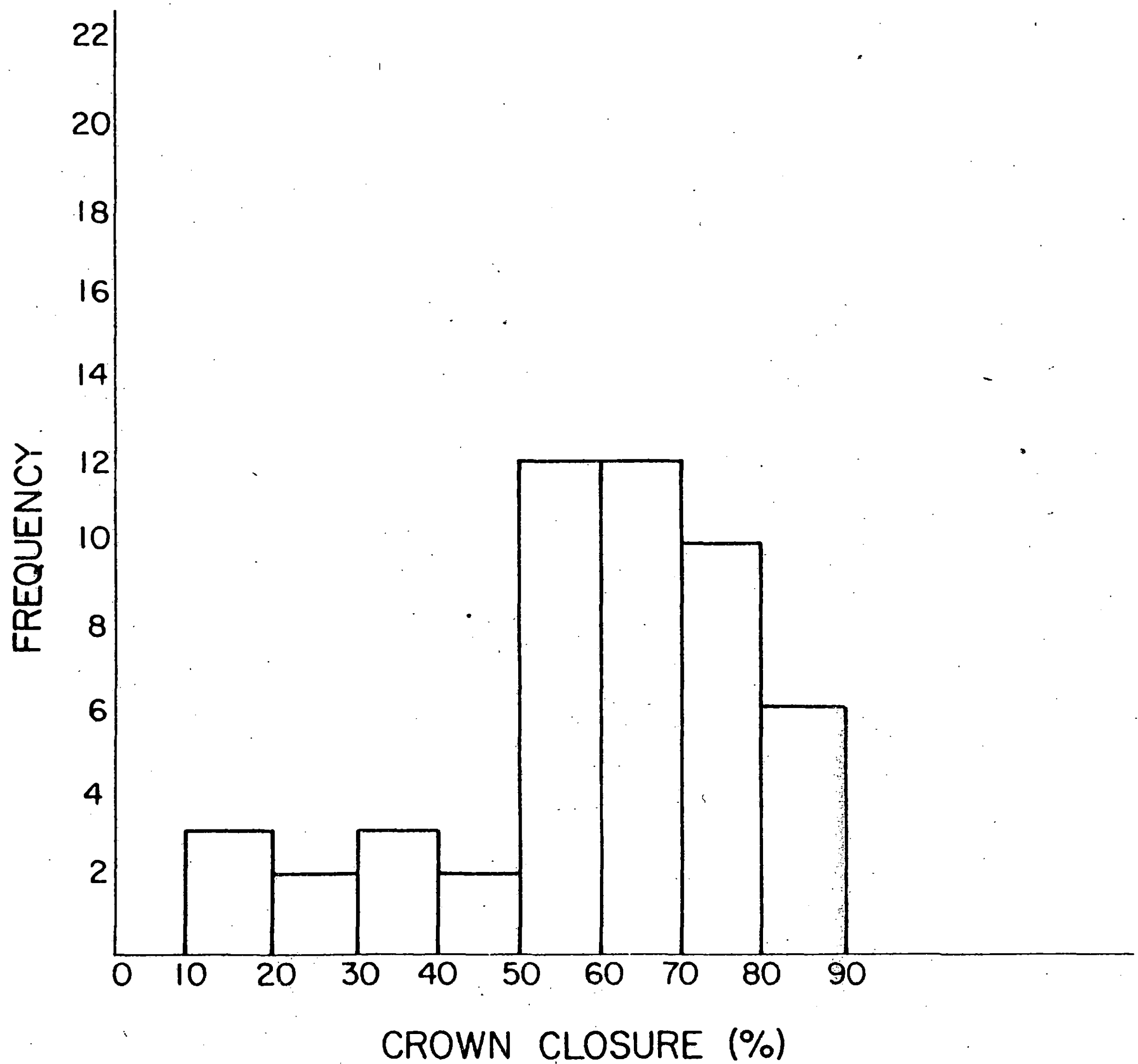

Figure 2. Frequency of crown closure classes for the 50 one-fifth acre volume table plots.

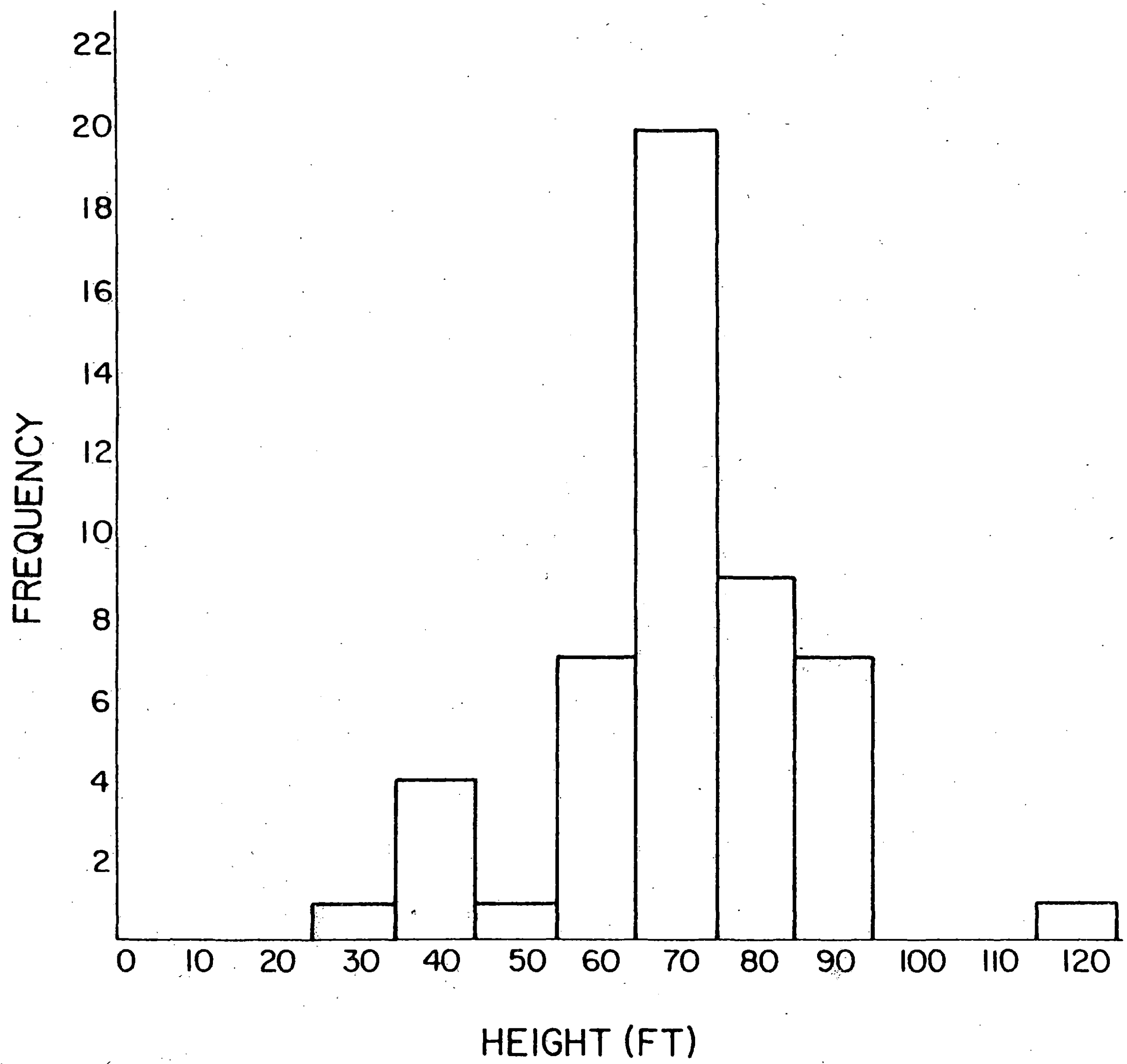

Figure 3. Frequency of height classes for the 50 one-fifth acre volume table plots.

Volume was defined as total volume (o.b.) per acre of all trees 6 inches in dbh and larger. Actual ground volume was estimated by inserting diameters and heights into an appropriate tree volume equation (Burkhart, 1977). These individual tree volumes were summed for each fifth-acre plot. These figures were then expanded to give volume per acre.

B. Preliminary Model Investigation:

The relationships between the possible predictor variables and volume, considering the mathematical restrictions on the predicted values generally indicated which prediction models were superior. Graphs were prepared to analyze variable trends, (i.e., volume vs. height and volume vs. crown closure vs. height). Figures 4 and 5 indicated a quadratic, curvilinear relationship between crown closure and volume, and between height and volume. Since volume is always positive (or 0), another logical model would predict the logarithm of volume from some form of crown closure and height. The use of a logarithmic transformation of volume would insure the non-negativity of the predictions. Thus, two families of models were explored. One family predicted untransformed volume per acre from the quadratic forms of crown closure and height. The other family of models predicted the natural logarithm of volume from various forms of closure and height.

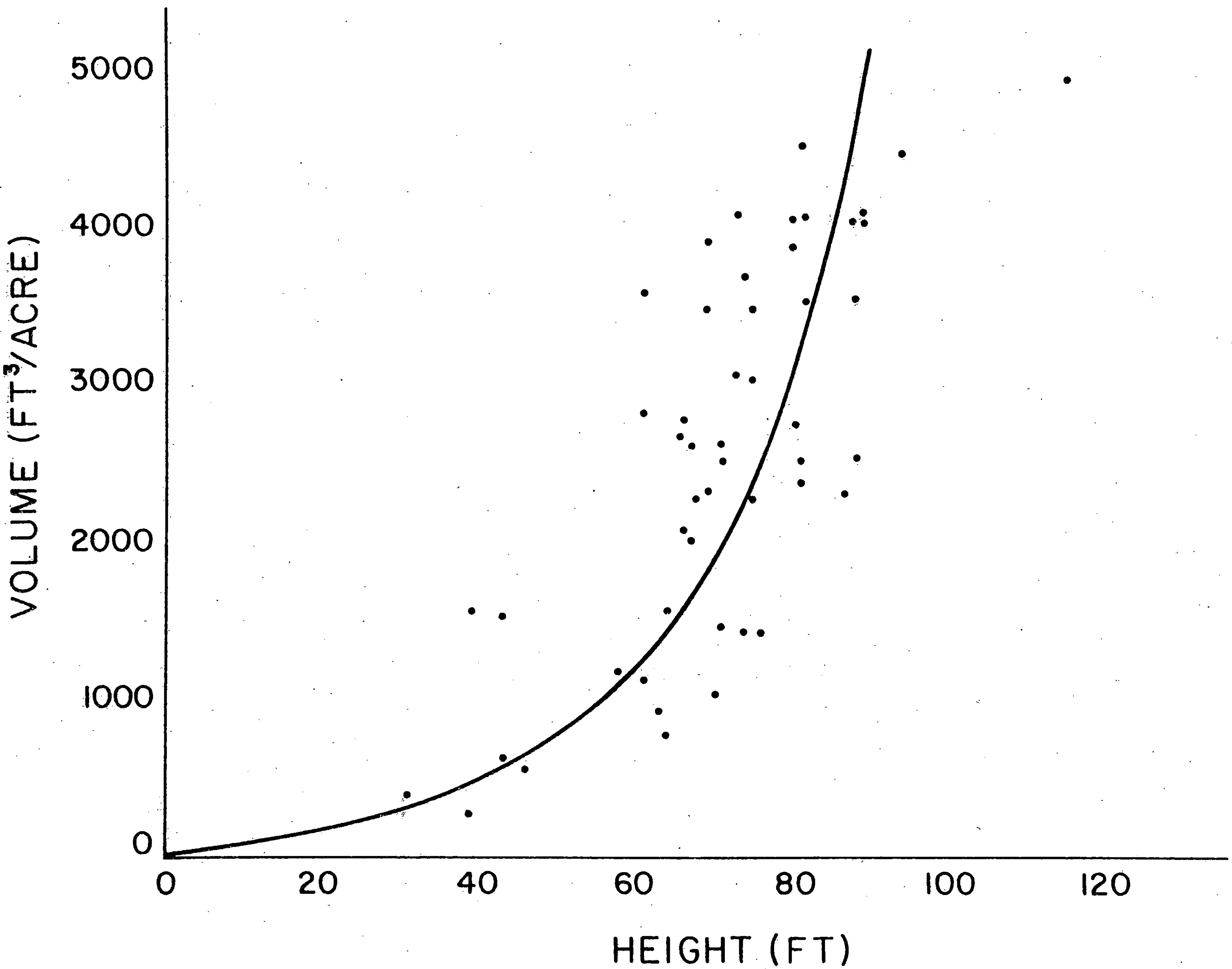

Figure 4. Plot of volume vs. height for the 50 plots with the hypothesized relationship indicated.

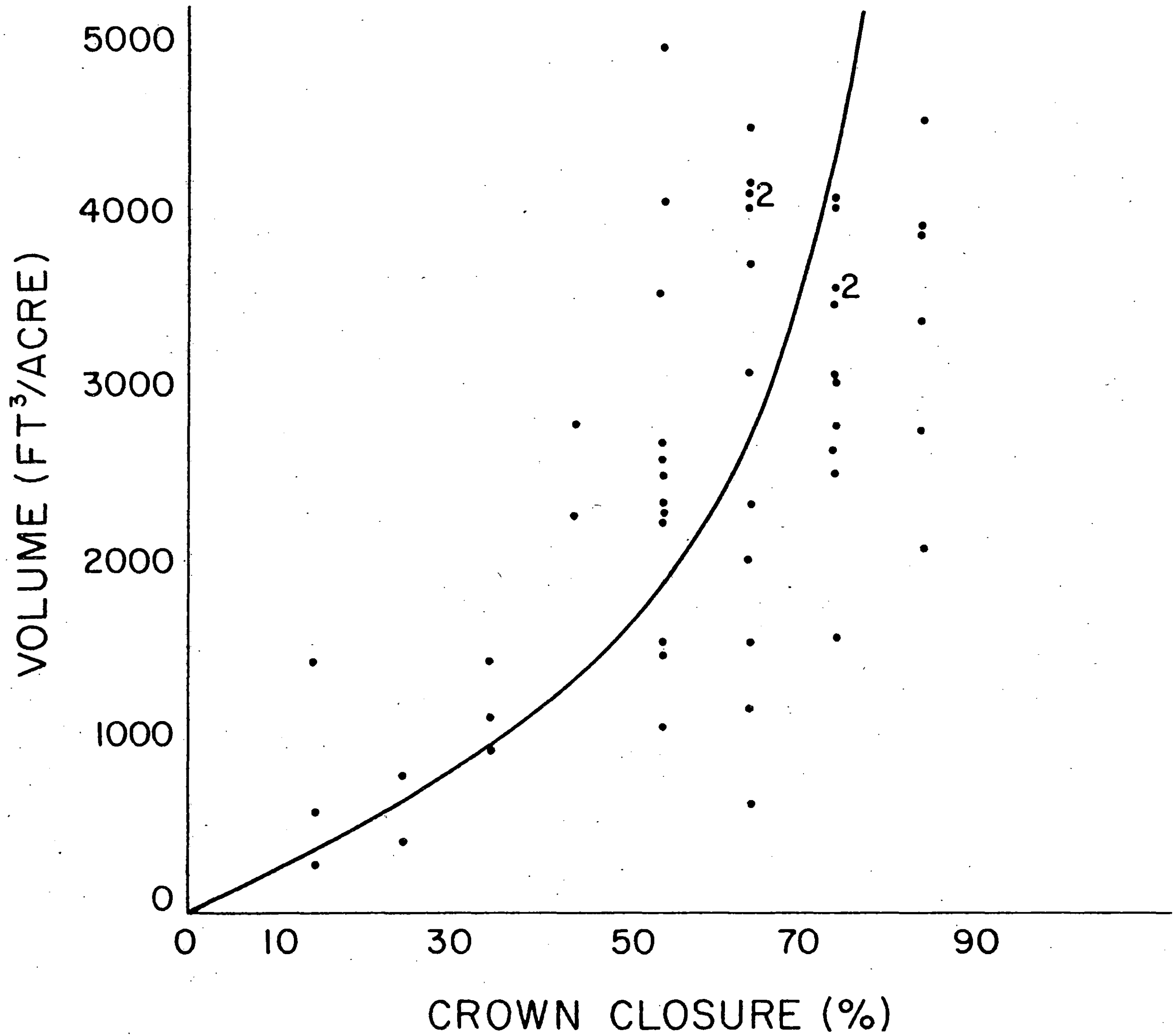

Figure 5. Plot of volume vs. crown closure for the 50 plots with the hypothesized relationship indicated.

C. Logarithmic Models (Model I):

The volume table data for crown closure, height and volume on the 50 one-fifth acre plots were fitted to many different logarithmic models using the Statistical Analysis System (SAS). These models were then compared on the basis of R^2, I, and the patterns of the plotted residuals. Although all the models had acceptable values for R^2 and I, they also exhibited systematic error trends for the lower range of crown closure (Figure 6). Since the model

$$\ln Vol = -2.12754900+0.76685628\ln CC+1.59065881\ln Ht^{*} \qquad \text{(Eq. 2)}$$

$$I = 620\ ft^3/ac$$

$$R^2 = 0.825$$

where Vol = volume (ft^3/acre)

CC = crown closure (%)

Ht = height (ft)

had the highest R^2 and lowest I, it was deemed the superior model fitted in the logarithmic family (see Table 1).

D. Quadratic Models (Model II):

Models which included various combinations of squared terms of crown closure and height were fitted to the volume table data using SAS. As forecased from Figures 4 and 5, the model with the superior fit (highest R^2, lowest I) was

---

*the symbol ln denotes the natural logarithm.

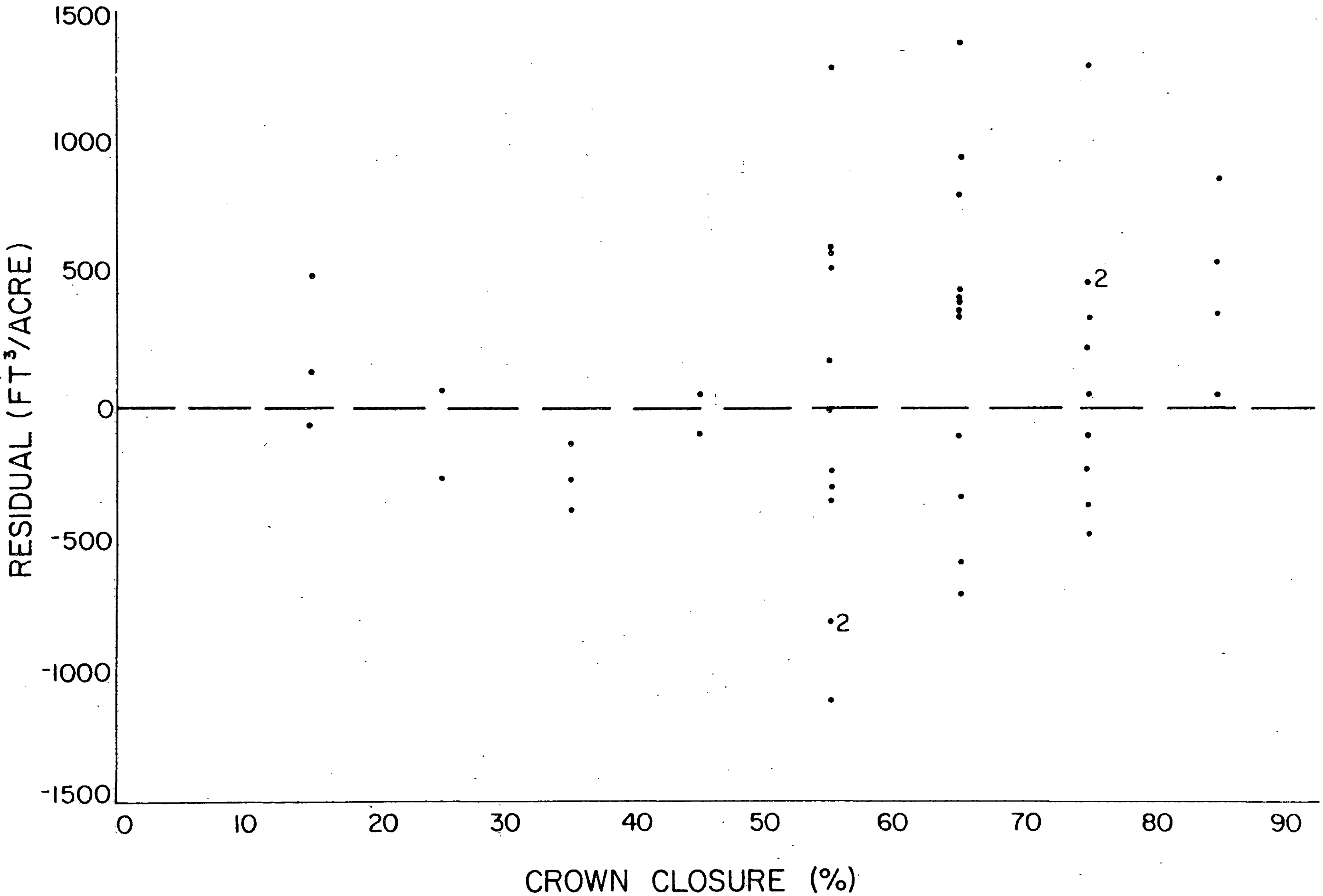

Figure 6. Plot of volume residuals over crown closure for Model I.

CROWN CLOSURE (%)	HEIGHT (FT) 20	30	40	50	60	70	80	90	100	110	120
5			140	210	280	350	440				
15			340	480	640	820	1010	1220			
25			500	710	950	1210	1500	1810			
35		410	640	920	1230	1570	1940	2340	2760		
45	260	490	780	1110	1490	1900	2350	2830	3350	3900	
55	300	580	910	1300	1730	2220	2740	3300	3910	4550	5220
65		650	1030	1470	1970	2520	3110	3760	4440	5170	
75		730	1150	1650	2200	2810	3480	4190	4960		
85			1270	1810	2420	3090	3830	4610			
95			1380	1970	2640	3370	4170	5030			

Table 1. Predicted (Model I) cubic foot volumes per acre for various values of crown closure and height.

$$Vol = -521.41615456+0.29112421(CC^2)+0.37170181(Ht^2) \quad \text{(Eq. 3)}$$

$$I = 620 \text{ ft}^3/\text{acre}$$

$$R^2 = 0.756.$$

Again, all the models exhibited systematic error trends in the lower range of crown closure (see Figure 7). However, predicting untransformed volume reduced this problem considerably (Table 2).

E. Final Model Selection:

In many cases, preliminary model investigation indicates a single "best"[1] prediction model. However, in this study, two dissimilar models seemed to fit the data well. Is one of these models superior to the other? The logarithmic model had a higher R^2 than the quadratic model, although both had acceptable values. Both models had values for Furnivals Index (I) of approximately 620 ft^3/acre (Furnival, 1961). However, a difference exists between the two models in the lower range of crown closure (25-45%). The residuals of the logarithmic model exhibited a marked curvilinear trend in the lower range of crown closure (Figure 6). Although the quadratic model exhibited somewhat the same trend, the systematic nature of the residuals was reduced when this model was used (Figure 7). The two models also differed in their respective negative volume prediction regions. Since one model predicted the natural logarithm of volume, no negative volume predictions were possible. However, the second model had a small region where

---

[1] best in the sense that it was the superior model investigated, not best among all possible models.

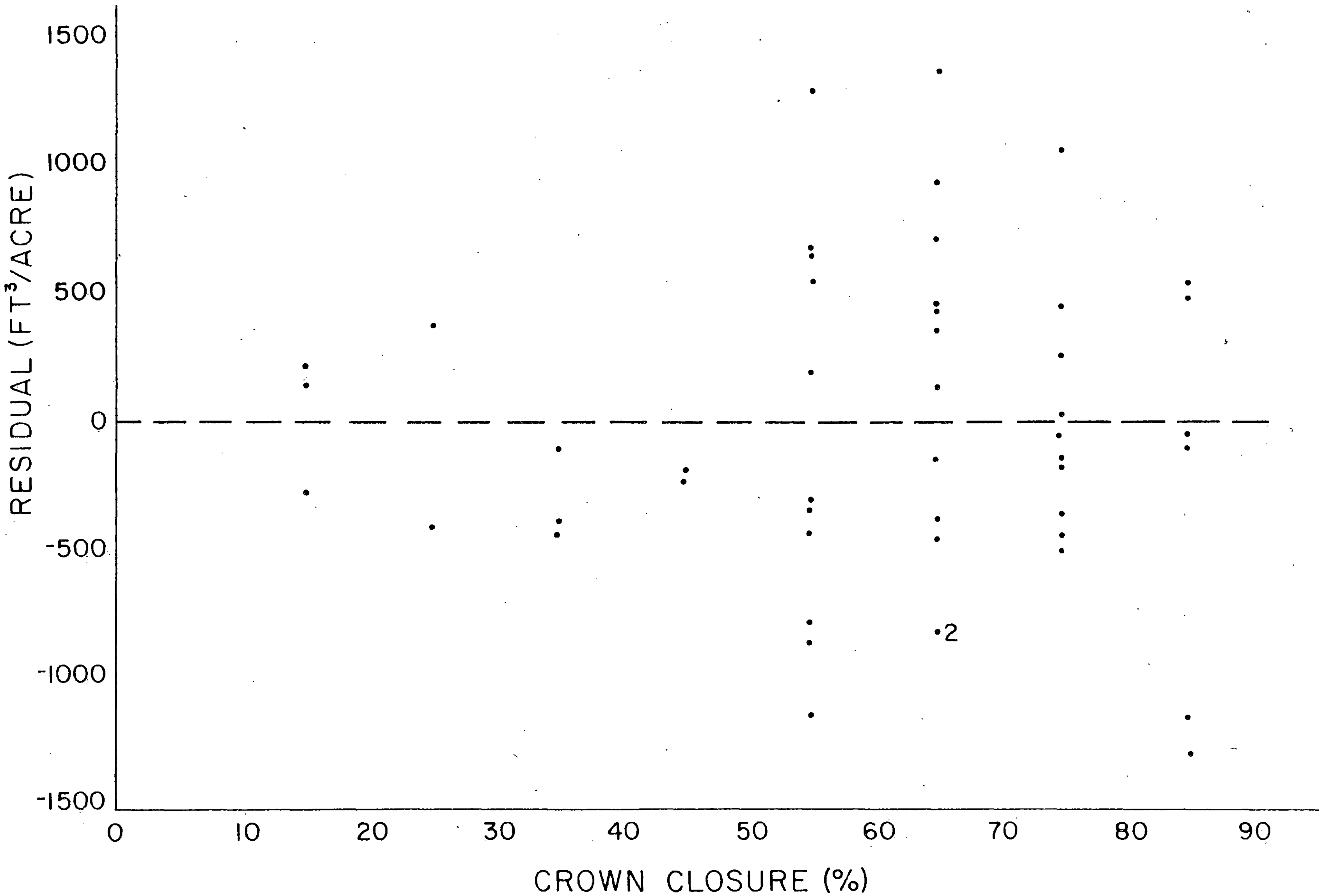

Figure 7. Plot of volume residuals over crown closure for Model II.

HEIGHT (FT)

CROWN CLOSURE (%)	20	30	40	50	60	70	80	90	100	110	120
5			80	420	820	1310	1860				
15			140	470	880	1370	1920	2550			
25			260	590	1000	1480	2040	2670			
35		170	430	760	1170	1660	2210	2850	3550		
45	220	400	660	1000	1400	1890	2450	3080	3790	4570	
55	510	690	950	1290	1700	2180	2740	3370	4080	4860	5710
65		1040	1300	1640	2050	2530	3090	3720	4430	5210	
75		1450	1710	2050	2450	2940	3490	4130	4830	5610	
85			2180	2510	2920	3400	3960	4590	5300		
95			2700	3040	3440	3930	4880				

Table 2. Predicted (Model II) cubic foot volumes per acre for various values of crown closure and height.

negative volumes were predicted (see Figure 8). A large area of 0 predictions was expected since we are predicting only the volume of trees in the 6-inch class and above. A stand with all trees less than 6 inches in dbh will have 0 volume, while still having positive values for crown closure and height. The logarithmic model predicts a positive volume for all values of crown closure and height. Thus, neither model seemed superior in all respects. A definite choice of one model was difficult and subject to personal bias.

F. Model Evaluations:

An independent evaluation of the regression model was accomplished by predicting the volume of five small areas attacked and killed by southern pine beetle. The mortality spots were located on the photographs, and their areas estimated using a micro-dot grid with 4225 dots per square inch. Crown closure was visually estimated by a single interpreter. Since the mortality spots had been ground checked, actual ground measurements of height and diameter* were available. The results of the evaluation are shown in Appendix A. There was little difference between the models in the range of this test, with the logarithmic model appearing slightly superior. However, it should be remembered that the range of the evaluation was considerably less than the range of data used to

---

*Again, Burkhart's tree volume equation was used to compute actual volume (Burkhart, 1977).

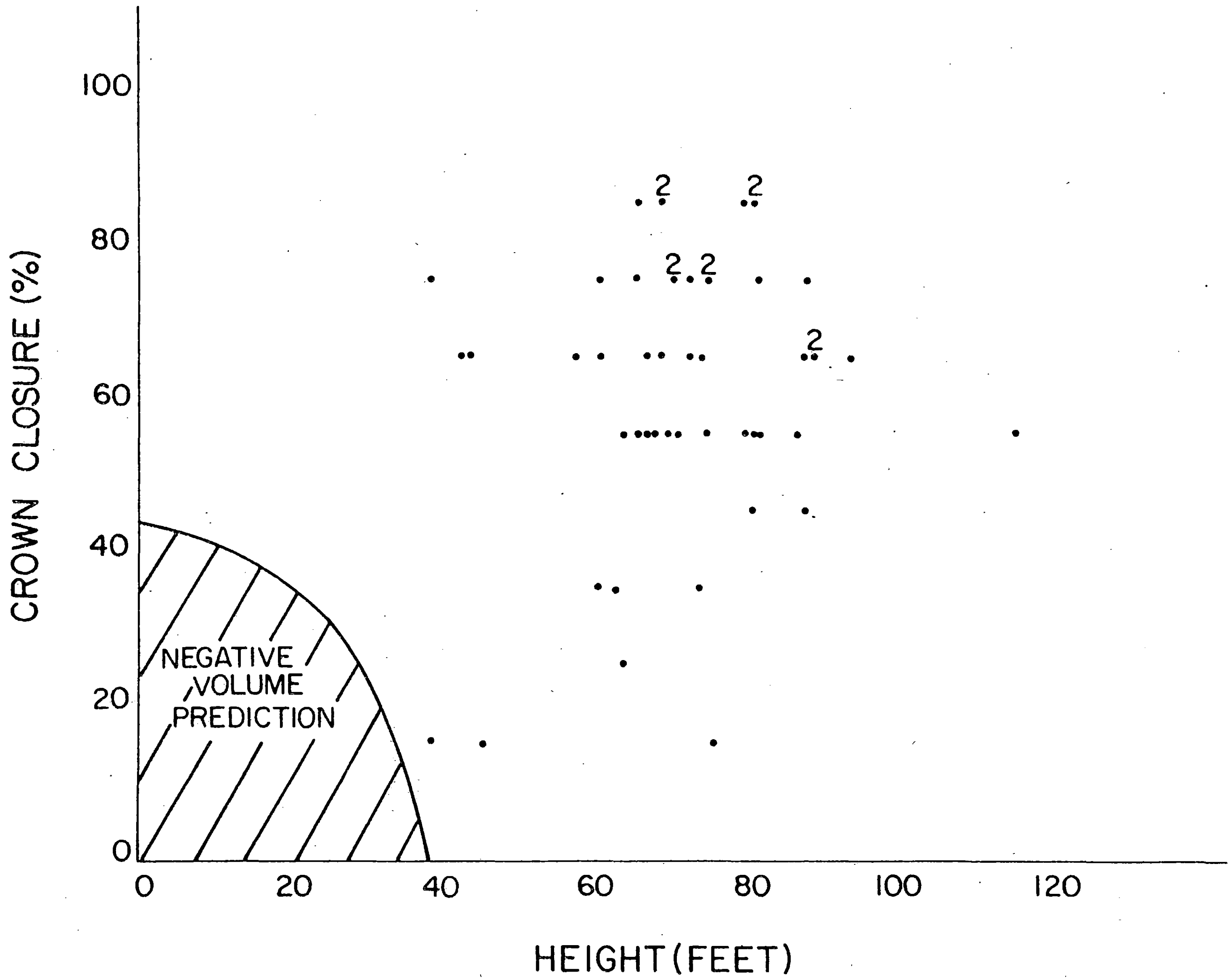

Figure 8. Graph of crown closure over height for the 50 one-fifth acre volume table plots, with the negative volume prediction region of Model II indicated.

onstruct the volume table. There were no extreme points checked and this is where the model will differ significantly. Also, note that spot volume estimation depended on estimating crown closure and area. Thus, this was less an evaluation of the regression model than an evaluation of the spot volume estimation procedure itself. The test did indicate that the method was satisfactory, but it offered little aid in deciding between the two models previously selected.

G. Comparison with Avery's Model:

One of the objectives of this study was to compare the newly constructed aerial photo volume table to Avery's Composite Aerial Photo Volume Table for Hardwoods and Pines in Northwest Mississippi (Avery, 1958). Since Avery's method for volume table construction was different from the one utilized in this study, only an empirical comparison of the models was possible. Graphs of volume vs. crown closure over 4 values of height were prepared (Figures 9, 10, 11 and 12). The plots clearly indicate the differences between the two new models and Avery's model. In general, the logarithmic and quadratic models predict higher volumes, and the difference increases as height increases. The volumes of the 5 southern pine beetle mortality spots (see Model Evaluations) were again predicted, but this time using Avery's model (Appendix A). The evaluation clearly indicated that Avery's volume model seriously underpredicted pine volume in this area. This was expected since the new models were constructed for pines only,

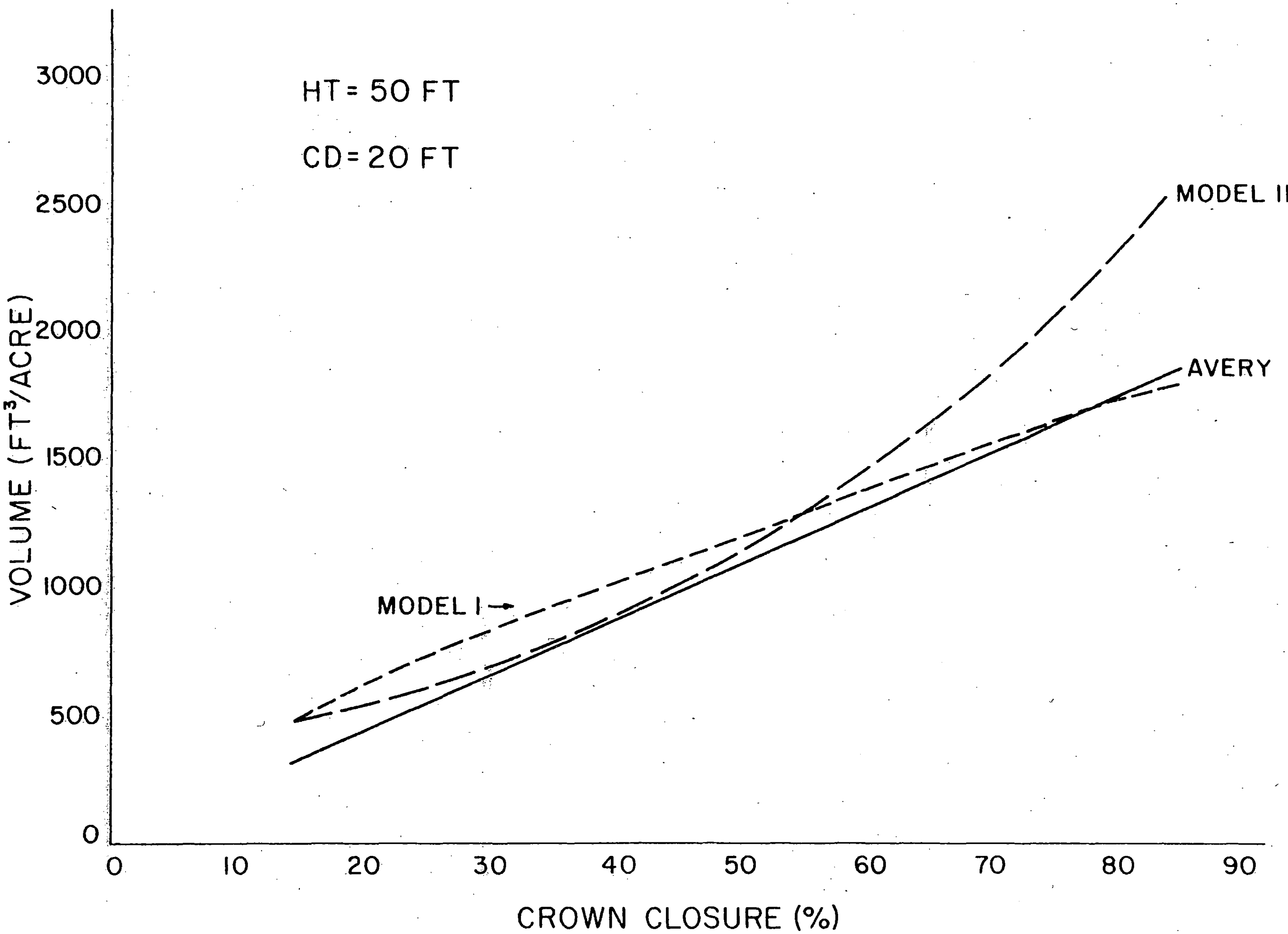

Figure 9. A graphical comparison of Avery's volume prediction model, the logarithmic model (Model I) and the quadratic model (Model II) when height is 50 feet and crown diameter is 20 feet.

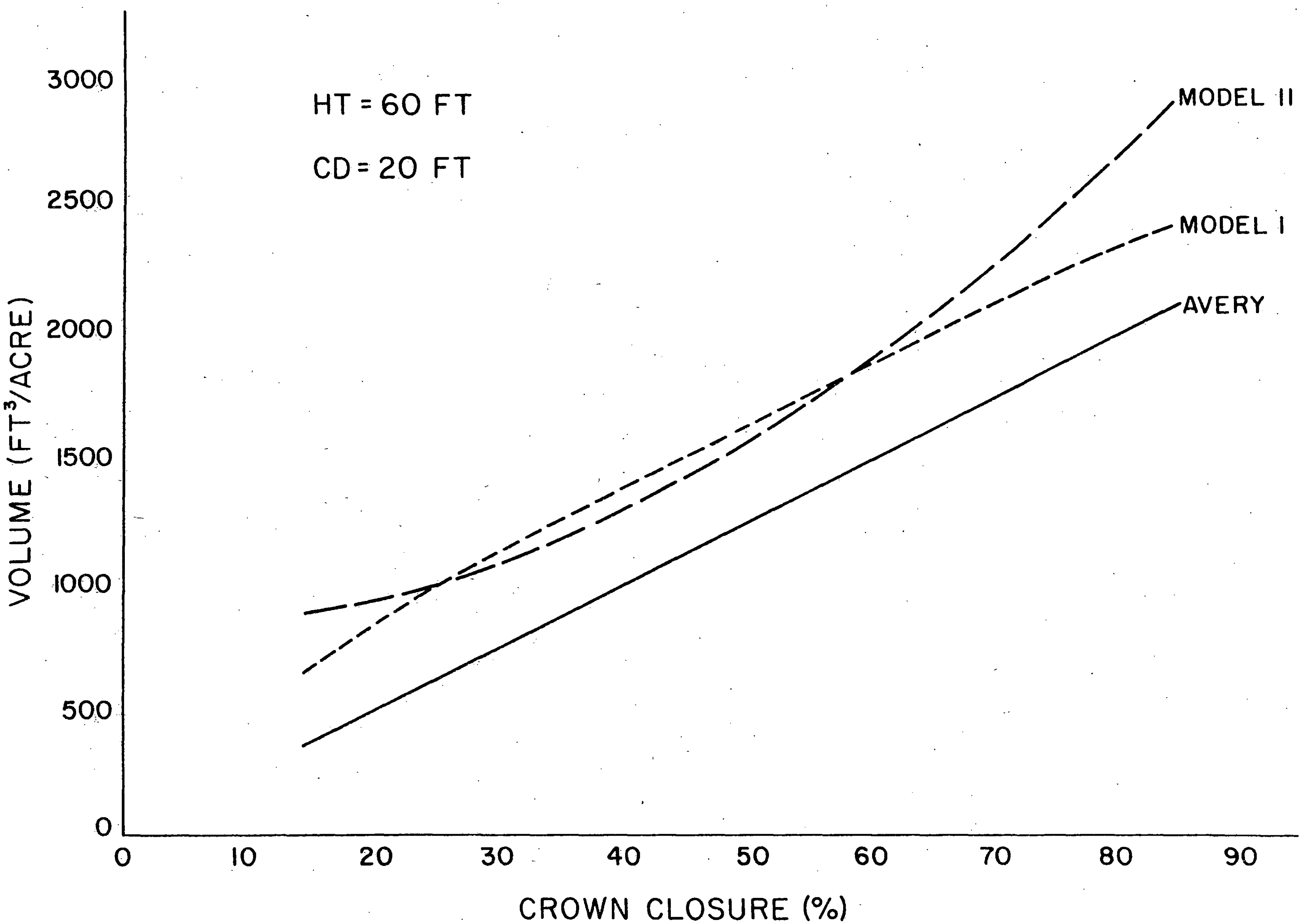

Figure 10. A graphical comparison of Avery's volume prediction model, the logarithmic model (Model I) and the quadratic model (Model II) when height is 60 feet and crown diameter is 20 feet.

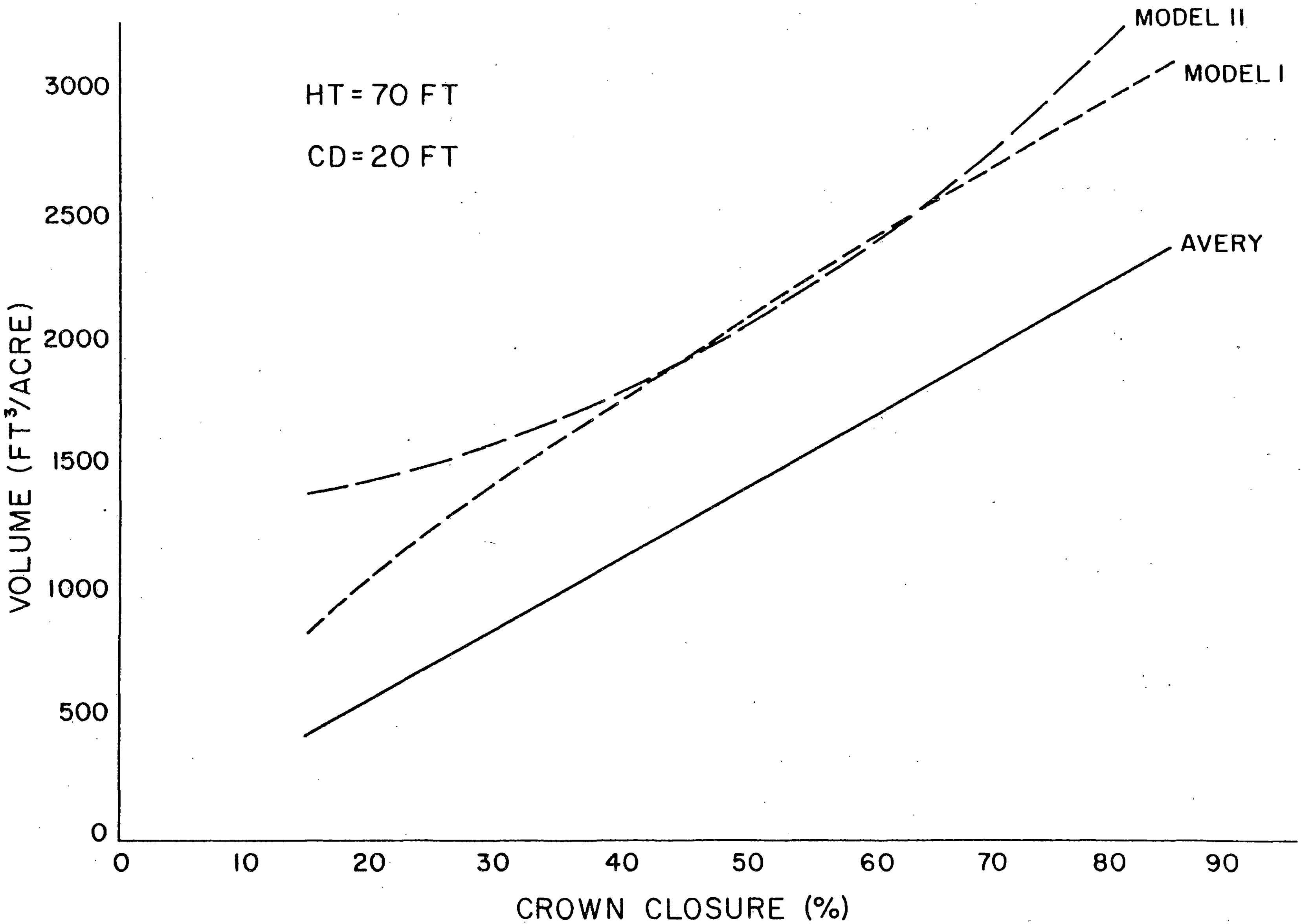

Figure 11. A graphical comparison of Avery's volume prediction model, the logarithmic model (Model I) and the quadratic model (Model II) when height is 70 feet and crown diameter is 20 feet.

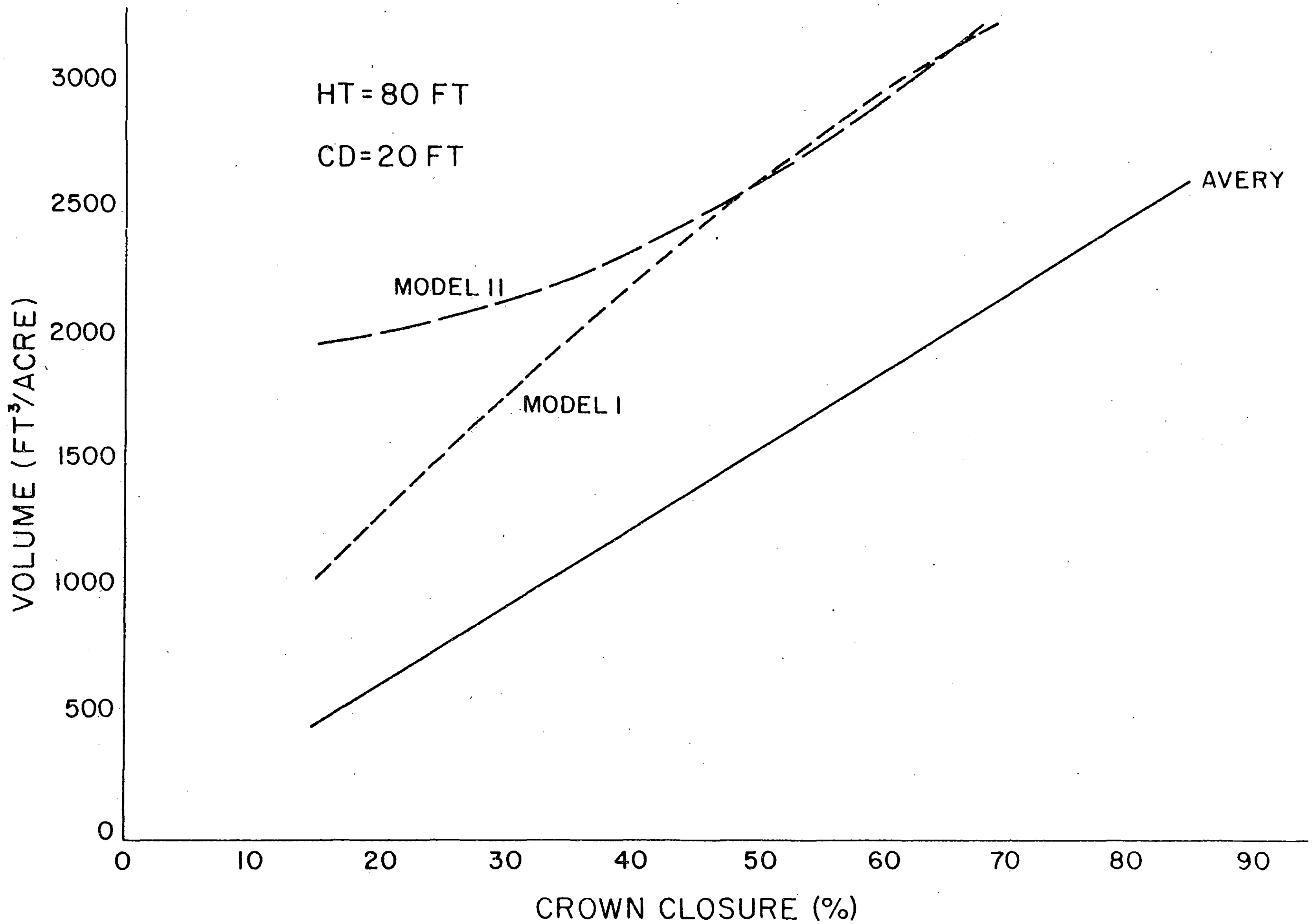

Figure 12. A graphical comparison of Avery's volume prediction model, the logarithmic model (Model I) and the quadratic model (Model II) when height is 80 feet and crown diameter is 20 feet.

not mixed stands. In summary, Avery's aerial photo volume table is not suitable for predicting the volumes of pine stands in Central Mississippi.

RECOMMENDATIONS AND FUTURE WORK

The model selection procedure resulted in two relatively equal models being selected. The decision to use a single model is thus dependent on personal judgment and intended use. Each model has inherent advantages and disadvantages. The logarithmic model has a higher R^2 and permits no negative volume predictions, but the model appears extremely biased in the lower range of crown closure. Consequently, the personal choice of the authors was the quadratic model. It had acceptable values for R^2 and I, while reducing the bias problem encountered in the lower range of crown closure. Also, any negative volume predictions could be reset to zero, which follows the logic of zero predictions put forth in a previous section. Whichever model is chosen by the user, the strengths and weaknesses of the model selected should be weighed and not forgotten.

Although this investigation is now complete, the possibilities for future work in this area are wide. Every aerial photo volume table should be updated when new applicable data becomes available. Also, plantations and natural stands are easily identified by photo interpreters. Thus, with proper data collection, a separate aerial photo volume table could be produced for the two types of stands.

Many advanced statistical techniques were not utilized in this study. Segmented polynomial regression, nonlinear estimation, the use of the PRESS statistic and biased estimation should be considered in any future endeavors in this area. Although nominally complete, this investigation, or offshoots from it, could continue on for many years and achieve meaningful results.

LITERATURE CITED

Avery, T. Eugene. 1958. Composite aerial volume table for southern pines and hardwoods. Journal of Forestry 56:10, p. 741-745.

Burkhart, Harold E. 1977. Cubic-foot volume of loblolly pine to any merchantable top limit. Southern Journal of Applied Forestry 1: 2, p. 7-9.

Furnival, G. M. 1961. An index for comparing equations used in constructing volume tables. Forest Science 7:4, p. 337-345.

Appendix A. Southern Pine Beetle Mortality Spot Evaluations for the Three Volume Prediction Models.

Spot	Height (feet)	Crown Closure (%)	Area (acres)	Actual Spot Volume (ft^3)	Model	Predicted Volume Per Acre (ft^3/acre)	Predicted Spot Volume (ft^3)	Residual
Sc-2-48	86	75	.3088	1234	I	3901	1204	+40
					II	3865	1194	+30
					Avery	2640	815	+419
Sc-2-15	76	65	.0320	150	I	2872	92	+58
					II	2855	91	+59
					Avery	2060	66	+84
Ne-2-11	83	65	.0446	260	I	3303	147	+113
					II	3269	146	+114
					Avery	2060	92	+168
Sc-4-62	83	85	.1606	473	I	4059	652	-179
					II	4142	665	-192
					Avery	2700	434	+39
Ra-5-30	88	55	.1646	579	I	3190	525	+54
					II	3238	533	+46
					Avery	1930	318	+261

Appendix B

Constructing a Multiple Linear Regression Volume Prediction Model

The utility of an aerial photo volume table cannot be denied. If properly constructed and used, it can provide a relatively precise estimate of volume per acre at much less expense than ground cruising.

Most aerial photo volume tables are just multiple linear regression models; a powerful and well understood statistical tool. The next few paragraphs will attempt to explain the step-by-step procedure for constructing a multiple linear regression volume prediction model. Enough elementary linear models theory is presented to ascertain the proper use and applicability of the model.

Data Collection

The most important single step in the construction of a good aerial photo volume table is the planning of, and the execution of data collection. Many things should be decided before any plane flies, or any foot touches the ground.

1. What do you want to predict? Is the variable of interest merchantable volume, total volume, inside bark or outside bark?
2. Where do I want to predict it? Is the population of interest clearly defined as to species and location?
3. What variables are available to predict it? Can I obtain measures of crown closure, height, crown diameter, number of trees per acre, etc.?

4. What scale and type of imagery would be best for this purpose? Should I use false color or natural color? What scale is most efficient?

5. How much, and what kind of data do I need? Is stratification planned? What will be the ground data collection procedures?

Of course, the answer to any of these questions affects the answers to the remainder, i.e, the chosen scale of the imagery certainly affects the measureability of the predictor variables.

Constructing A Multiple Linear Regression Model

The objective of an aerial photo volume table is to quickly and cheaply estimate the volume of a stand of timber using only photo measureable variables. Logically, linear regression is a tool to use to achieve this end. Linear regression is a mathematical method whereby the expected value of one variable can be predicted from the values of other variables correlated with it. We must assume that there is a linear[1] relationship between the predictor (independent) variables and the variable of interest (volume, in our case). Depending on your objectives, other assumptions may also be

---

[1]Linear in the parameters, i.e., a model with a parameter in the exponent is nonlinear, but a model containing logarithms of independent or dependent variables is linear.

Nonlinear - $e^{\beta_o} DBH^{\beta_1} HT^{\beta_2} e^{\varepsilon}$

Linear - $\beta_o + \beta_1 \ln DBH + \beta_2 \ln HT + \varepsilon$

applicable, but this is beyond the scope of this short treatise.

The first step in constructing the regression model would be to examine the data for functional relationships between the possible predictor variables and the dependent (predicted) variable. Note here that transformations of the independent variables (logarithms, squares, etc.) are valid predictors. Also available are interaction terms, i.e., combinations[2] of the various independent variables. Many times the variables themselves are poor predictors, while transformations of them and/or interaction terms are more than adequate. Since volume itself is always positive (or 0), any transformation of volume which restricts predictions to the positive numbers should be included in the analysis. To illustrate the many model possibilities, assume you wish to estimate volume (V) from crown closure (CC) and height (HT). You might predict V or $\log_e V$ from CC or HT or CC^2 or HT^2 or $\log_e$ CC or $\log_e$ HT or CC/HT or HT/CC or HT·CC or more, or any combination of these. If all avenues are explored, you will either produce a good model, or conclude that a linear relationship between the variables cannot be modeled. Data examination usually consists of examining the graphs of the dependent variables over the various independent variables, and of comparing partial correlation coefficients. Are any relationships indicated? Can any variables be eliminated? Considering the complexity of the problem, any hints which help you reduce the confusion should be welcome and considered.

---

[2]Cross products, divisions, etc.

A set of possible models has now been identified. The next step is to compare these many models and to make a decision concerning which one you feel is superior. Comparisons are generally made on the basis of R^2, $\hat{\sigma}$ and the patterns of the plotted residuals. The coefficient of determination, R^2, is defined as the percent of the total variation explained by the regression model. Thus, the higher the R^2, the better your model fits the data set. However, it doesn't necessarily follow that the model with the highest R^2 is the best volume predictor for the population. The model with the lowest $\hat{\sigma}$ (also called $Sy \cdot x$, Mean Square Error - an estimate of the variance around the regression) does give the most precise prediction, however. If the dependent variable has been transformed, say your model predicts the logarithm of volume, then the computed $\hat{\sigma}$ is the precision of predictions of the logarithm of volume, not volume itself. The value for $\hat{\sigma}$ must be transformed back into the original volume units using a process set forth by Furnival (1961) if comparisons are to be made between untransformed and transformed made. Another important comparison tool is the model residuals. Residuals are defined as the observed value of the dependent variable minus the value for the dependent variable predicted by the regression model. A residual is computed for each point in the data set, and these values are plotted over each independent variable. If the model is unbiased, the residuals should be randomly scattered around a horizontal line through the residual value of 0. This would mean

that you are as likely to overpredict as underpredict across the range of the independent variables. If this is true, as the number of times you use the model increases, the average of the predictions will approach the true mean of the population at each X (independent variable). Thus, if you wish an unbiased model[3], this is an important comparison tool. You now have three ways to compare the various models. Of course, there are many more advanced comparison techniques, but those given in this paper should suffice in this simple situation.

Rarely will any single model be superior in all respects, thus, the final model decision is somewhat colored by personal judgment. Perhaps the final step in the model process is to ask the question: Does the model I chose make sense? Does the final model fit in with the known biologic processes inherent in the problem? If the answer was "no," the next question should be: Do I care that the model seems purely artificial?[4] Is the model unnecessarily complex? It is evident that the user must make the final decision. Statistics will not do it for you, only aid you.

Evaluation of the Model

The final step in any volume estimation procedure should be an independent evaluation of the predictive power of the model.

---

[3] It is not uncommon in statistics to use biased models. They are sometimes superior to unbiased models in other ways, such as the magnitude of the variance.

[4] Never assume that a cause and effect relationship exists between the dependent and independent variables. A linear regression model is a mathematical abstraction, possibly based upon a cause and effect relationship.

Up to this point, all model comparisons were based on indicators of data set fit, not indicators of how well the model predicts mean volume. Thus, the need for an independent evaluation of the chosen model is clear. Also, if a single model could not be deemed superior following previous comparisons, the model evaluation step could be used as the final comparison tool, i.e., select the model that was the "best" predictor on the new data set.

Evaluating the model is actually easier than constructing it. Ths independent[5] data set should contain all the independent variables required by the models, coupled with an actual measure of the variable to be predicted. All the variables should be measured exactly as they were for the original data set, and they should span the range of the original data. The independent variables are inserted into the model, and a predicted value for the dependent variable is produced. A residual is formed for each data point by subtracting the predicted value from the observed value. Examine these residuals. Are they very large? Does there seem to be any bias in the model, i.e., does the model tend to overpredict or underpredict in some range of the independent variables? Again, your judgment is required for the final decision. If the evaluation is considered negative, select the model judged

[5] The new data set should contain 1/3 - 1/2 as many points as the original data set.

next best and repeat the evaluation procedure until a satisfactory model is found.

Use of the Final Model

There are many things to remember about using regression models. First, and most important, do not try to predict outside the range of the data used to construct the model. The behavior of the model cannot be guaranteed outside the appropriate range. Second, always use the regression equation, not the table. The table is an illustrative tool, used only to present certain outcomes of the model in tabular form. Three, the predictor variables should always be measured in the same way as the original variables were. If differences exist, they must be accounted for in some way. Fourth, regression models predict only the mean response at each level of the independent variables. The models are meant to be used repeatedly. Lastly, always remember that estimation produces error. You minimize that error by using reliable data to construct the model, and by using reliable data in the model.